PAYMENT FORM DISCRIMINATION METHOD AND APPARATUS

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to a payment form discrimination method and apparatus for discriminating payment forms for paying taxes or public utilities fees at a financial institution or the like, and in particular to a payment form discrimination method and apparatus that searches for a character string by which forms can be discriminated and discriminates the forms by the search results.

2. Description of the Related Art

The payment of fees, such as taxes or public utilities fees, can be carried out at a place that is convenient for the person making the payment, like a financial institution, a post office, a convenience store, etc. For such a payment of fees, a payment form stating the payment amount is sent to the person who has to make the payment. The payer then takes the payment form to a financial institution or the like, where the payer can pay the fees. There is a strong demand for the automation of such a fee payment process.

Usually, methods for automatically recognizing

characters on the payment form are used for automatic payment processing. For this automatic recognition, there is no need to recognize all characters on the payment form. For example, it is sufficient to recognize the payee, the payment amount and the payment time limit on a tax payment form, and there is no necessity to recognize captions or the like. In order to recognize the characters of the relevant portions among the multitude of characters on such a payment form, it is advantageous to register formats of the form in advance, cut out the registered regions from the read in image of the form, and perform character recognition with the registered character types.

Moreover, since there are several types of forms with different formats that can be processed automatically, it is necessary to identify the format type, access the identified format of the form, and carry out a character recognition depending on the form type. The following techniques have been proposed as automatic identification methods of the form type:

(1) Subjecting a group of characters indicating the form type, such as a title on the form (for example, payment notification slip) to automatic recognition, so as to identify the form type.

(2) Extracting rules from the form and identifying the rule type by the format of the rules.

(3) Attaching form discrimination information, such as a barcode, on the form to be recognized, and reading this

information to identify the form.

However, these techniques pose the following problems:

(1) The method of recognizing a form title poses the problems that the recognition process takes a long time since the title is in Chinese (Kanji) characters, and there is the possibility of confusion with forms of similar titles. For example, in payment forms for municipal taxes, similar forms may be used for city A and for city B, so that it is not possible to recognize the forms by the title "Local Taxes".

(2) The recognition with rules takes a long time to process, and there is the problem that when the rule formats are similar, the form may easily be identified incorrectly.

(3) With the method of attaching form discrimination information, the forms can be discriminated reliably, but there is the problem that only those forms can be identified to which discrimination information has been attached. That is to say, it is not possible to identify those forms currently in circulation to which such discrimination information has not been attached.

SUMMARY OF THE INVENTION

Consequently, it is an object of the present invention to provide a payment form discrimination method and apparatus, with which payment forms can be accurately

identified even when no discrimination information is attached to them.

It is another object of the present invention to provide a payment form discrimination method and apparatus, with which payment forms can be identified quickly.

It is yet another object of the present invention to provide a payment form discrimination method and apparatus, with which payment forms can be identified quickly using one portion of the recognition data on the payment form.

In order to attain these objects, a payment form discrimination method of the present invention for discriminating payment forms which state at least a payee account number and a payment amount, includes a step of acquiring an image of the form; a step of searching the payee account number in the image in accordance with an account number searching rule; and a step of discriminating the type of form based on the searched payee account number.

A payment form recognition apparatus of the present invention includes an image reader for acquiring an image of the form; and a processing unit, which makes a search for the payee account number in the image in accordance with an account number searching rule, and discriminates the type of form based on the searched payee account number; wherein the processing unit recognizes respective areas of the image by corresponding recognition categories, depending on the type of form that has been discriminated.

The present invention takes advantage of the fact that the payment form contains the payee account number, and that the form type is specified by this account number, so that the form is identified by recognizing the payee account number from the form image. Moreover, the account number is a character string of numerals and symbols, so that the account number can be identified on the form image with high precision by simple search rules (knowledge information). Therefore, the form can be identified accurately and fast. Furthermore, the payee account number is a portion of the input data of the payee form, so that form identification and data input can be performed simultaneously.

When the searching step in the payment form discrimination method of the present invention includes a step of recognizing an account number with the searching rule; and a step of determining whether this recognized account number is a registered account number. Therefore, the account number can be recognized with a simple searching rule, while incorrect recognition results can be prevented by referencing the registered data.

When the searching step in the payment form discrimination method of the present invention is a step of making a search for the payee account number in accordance with a searching rule that is based on recognition categories and regularities regarding a character string, then the account number can be searched with a simple rule,

and the process speed can be increased.

When the searching step in the payment form discrimination method of the present invention includes a step of recognizing a plurality of account numbers on the payment form, and a step of merging a plurality of result which have been recognized to determine the payee account number, then incorrect recognition results can be prevented.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram of a form recognition apparatus in an embodiment of the present invention.

Fig. 2 is a diagram illustrating a payment form to be recognized by the apparatus in Fig. 1.

Fig. 3 is a diagram illustrating a specific example of a tax payment form to be recognized by the apparatus in Fig. 1.

Fig. 4 is a diagram illustrating a specific example of another tax payment form to be recognized by the apparatus in Fig. 1.

Fig. 5 is a diagram illustrating a public utilities fees payment form to be recognized by the apparatus in Fig. 1.

Fig. 6 is a first part of a flow chart of a form recognition process in an embodiment of the present invention.

Fig. 7 is a second part of a flow chart of a form recognition process in an embodiment of the present invention.

Fig. 8 is a third part of a flow chart of a form recognition process in an embodiment of the present invention.

Fig. 9 is a diagram illustrating the operation of the form recognition process in Fig. 7.

Fig. 10 is a flowchart of the form registration process in an embodiment of the present invention.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The following is an explanation of the present invention, divided into a form recognition apparatus, a form discrimination process, and other embodiments of the present invention.

Form Recognition Apparatus

Fig. 1 is a block diagram of a form recognition apparatus, Fig. 2 is a diagram illustrating a payment form to be recognized, and Figs. 3 to 5 are diagrams illustrating specific examples of a payment form.

As shown in Fig. 1, an image scanner 1 includes a scan bed 10 on which the form 4 to be scanned is placed, and an optical scanner 11, which reads in an image of the form placed on the scan bed 10 and sends the image data

to a processing device 2. This image scanner 1 is of the non-contact type, and the scanner 11 is made of color CCD (Charge Coupled Device).

The processing device 2 is made of a desktop personal computer, and includes a data processing unit (CPU) 20, a display device 21, and files 22. The data processing unit 20 is made of a CPU, a memory, etc., and carries out a form recognition process. Before explaining the configuration of the files 22, the payment form 4 is explained, with reference to Figs. 2 to 5.

Fig. 2 is a diagram of the payment form 4, and shows a local tax payment form described in Chinese character. The city to which the taxpayer belongs sends this payment form 4 to the taxpayer, and the taxpayer presents this payment form 4 and cash at a tax payment counter of a financial institution or the like. The payment form 4 consists of a local tax receipt 40, a local tax payment slip 41, and a local tax payment notification slip 42. The receipt 40 is handed to the taxpayer at the tax payment counter. The payment slip 41 is kept by the financial institution or the post office as the payment counter. The payment notification slip 42 is kept by the city collecting the tax.

A payee account number 50 and a payment amount 51 are described in the receipt 40. The payee account number 50 is the city's account number for postal transfer, and the payment amount 51 is the amount of taxes paid.

Similarly, a payee account number 52 and 54 and a payment amount 53 and 55 are described in the payment slip 41 and the notification slip 42. A payment time limit 56 is further described in the notification slip 42. These are the elements to be recognized. As will be explained below, the account number is made up of numbers and hyphens, namely a 5-digit number, followed by a hyphen, another 1-digit number, a hyphen and a six-digit number. That is to say, it consists of fixed recognition categories and regular character strings.

Figs. 3 and 4 show specific examples of a payment 4 for local tax described in Kanji characters. Fig. 3 shows a payment form for local (municipal and prefecture) tax of *AAA* prefecture and *BBB* city. Fig. 4 shows a payment form for local (municipal and prefecture) tax of *CC* prefecture and *DD* city. Both consist of a payment slip 43 and a payment notification slip 44, into which the payee account number 50, the payment amount 51 and the payment time limit 56 are described. These are the elements to be recognized, whereas all other entries are reference information. Also in this form 4, the account number is made up of numbers and hyphens, namely a 5-digit number, followed by a hyphen, another 1-digit number, a hyphen and a six-digit number. That is to say, it consists of fixed recognition categories and regular character strings.

Fig. 5 shows a specific example for a payment form 4 for the collection of water services fees by *FF* prefecture,

which consists of a notification slip and receipt 45 and a payment notification slip 44, into which the payee account number 50, the payment amount 51 and the payment time limit 56 are entered. These are the elements to be recognized, whereas all other entries are reference information. Also in this form 4, the account number is made up of numbers and hyphens, namely a 5-digit number, followed by a hyphen, another 1-digit number, a hyphen and a six-digit number. That is to say, it consists of fixed recognition categories and regular character strings.

In these forms, the elements to be recognized are the same, but their position is different. In accordance with the present invention, as the payee account number 50 is specified to the payee issuing the payment form 4, and in order to discriminate between these forms, the payee account number 50 is recognized to identify the form.

Returning to Fig. 1, the file 22 has a search rule table 30, an account master table 31, and a form data table 32. The search rule table 30 stores the rules for searching the account numbers, and stores the rule by which the account numbers are configured and the account number identification categories. In the case of the above-described forms, the rule by which the account numbers are configured is that it is a numerical string within 5 digits - (hyphen) - 1 digit- (hyphen) - 6 digits, and the account number identification categories are numerals and symbols (hyphens). Based on this search rule,

the CPU 20 searches the image of the form for an account number, which is then recognized. Based on this search rule, the account number can be found and recognized on the image, regardless where it is located on the form image.

The account master file 31 is a master file of actual account numbers and is used to check whether a recognized account number actually exists. The form data table 32 is a table in which the formats of the various forms are stored, such as form IDs, form names, form data, account numbers, etc. The form data include, for example, the position of the characters to be recognized in the image and the recognition categories.

In an example of the present invention, account numbers are added into a conventional form data table 32, in addition to the search rule table 30 and the account master table 31. Thus, it is possible to specify the forms with the recognized account number, and character recognition using the form data in the table 32 becomes possible.

This form recognition apparatus can be installed at designated payment counters at financial institutions, post offices, convenience stores, or the like. The form recognition apparatus is connected to a host computer (not shown in the drawings).

Form Identification Process

Figs. 6 to 8 are flow charts of the form recognition

process in an embodiment of the present invention, and Fig. 9 is a diagram for illustration of this operation. Figs. 6 to 8 show the form recognition process carried out by the CPU 20 of Fig. 1.

S100: After the CPU 20 has acquired a binary image from the image reader 1 and stored it into its internal memory, it retrieves the recognition categories and the configuration rules from the search rule table 30 and starts to search the binary image.

S110: The CPU 20 determines whether the search of the binary image of the image reader 1 has been finished. If the search has been finished, the process advances to Step S160 in Fig. 7.

S120: Recognizing a specified recognition category (in this case, a hyphen) in the image data of the binary image, the CPU 20 searches for characteristic hyphens, as defined by the configuration rules. According to the configuration rules, characteristic hyphens are two hyphens arranged in series.

S130: When characteristic hyphens have been recognized, the surroundings of the hyphens are subjected to character recognition in the categories numerals and symbols (hyphens), which are the specified recognition categories.

S140: Then, it is decided whether the character string that has been subjected to character recognition is a numerical string within the configuration rules of

the account number, namely 5 digits - 1 digit - 6 digits. When the recognized character string does not match the configuration rules, the process returns to Step S110.

S150: When the recognized numerical string does match the configuration rules, then this numerical string is taken as the account number, thereby obtaining the recognition result. Moreover, a number extraction counter is updated. Then, the process returns to Step S110.

S160: After finishing the search of the form image in Fig. 6, the process advances to Fig. 7, and the CPU 20 decides whether a recognition result has been obtained. When no recognition result has been obtained, then the image search was unsuccessful, and the process advances to Step S230 in Fig. 8.

S170: When a recognition result has been obtained, then it is determined with the number extraction counter, whether there is more than one recognition result. If there is not more than one recognition result, that is, if there is only one recognition result, then the process advances to Step S190.

S180: When more than one recognition result has been obtained, then a process of merging the recognition results is performed. That is to say, the number of rejected characters in the recognition results is calculated. The number of rejections in recognitions results ①, ② and ③ for 5 digits - 1 digit - 6 digits is calculated as shown in Fig. 9. Here, "rejected character" means that it has

been determined that a character is a number, but it could not be determined which number. Then the character strings of the rejection results are merged. As shown in Fig. 9, when the number of rejections is not identical, then the merged result of a digit is taken to be the numerical character that has been recognized most often for that digit. When the number of rejections is identical, then the merged result of a digit is taken to be the numerical character of the recognition data with the fewer rejections. This is repeated for the each character in the character string, and the procedure advances to Step S190.

S190: If there is only one recognition result in Step S170, then this recognition result is confirmed as the recognition result, and if there is more than one recognition result and a merge process was performed in Step S180, then the merge result is confirmed as the recognition result.

S200: The process advances to Fig. 8, and it is determined by referencing the master file 31 whether the confirmed recognition result (account number) actually exists. If the account number does not actually exist, then the recognition has failed, and the process advances to Step S230.

S210: If the account number actually exists, then the form data table 32 is referenced with this account number, and it is determined whether a corresponding form exists. If there is no corresponding form, then the form cannot

be confirmed, and the procedure advances to Step S230.

S220: When the corresponding form has been confirmed, then, as conventionally, the character recognition data (positions of elements to be recognized and recognition categories) of the corresponding recognition portions are retrieved from the form data, and the data at the relevant positions of the form image data are subjected to a character recognition with these recognition categories. This is repeated for all characters to be recognized as specified by the form data, and the recognition results are sent to the host. Then, the form recognition process is finished.

S230: If the account number cannot be confirmed, then the form identification is carried out with another method. Here, an attempt is made to identify the forms by conventional rule extraction. That is to say, rule data are extracted from the binary image, and the form data in the form data table 32 are searched to find a form with matching rule data. When a form with matching rule data has been found, then the process advances to Step S220, and character recognition is performed.

S240: If no form with matching rule data can be found, then it is concluded that the form is a new form, the image is stored for form registration at a later stage (explained below with reference to Fig. 10), and the process is finished.

Thus, the payee account number described into the payment form is searched to identify the form, so that the

account number can be identified from the form image with high precision by simple search rules (knowledge information), because the account number is a character string of numerals and symbols. Therefore, the form can be identified accurately and fast. Furthermore, since the payee account number is a portion of the data input into the payment form, the form identification and the data input can be carried out simultaneously.

The hyphens are detected on the form image, and their surroundings are subjected to character recognition, so that the account number can be recognized with high speed and precision, even when the position of the account number is not specified. Furthermore, due to the merging process, incorrect recognition results can be discarded. Similarly, incorrect recognition results can be completely prevented by reference to registered account numbers in a master file.

Fig. 10 is a flowchart of the file registration process in an embodiment of the present invention.

S300: The image data, form data file 32 and account number master file 31 of a file 22 are entered into a memory by operating a keyboard (not shown in the drawings) connected to the CPU 20.

S310: The entered image data are displayed on a display device 21.

S320: An operator observes the display, defines the recognition portions (positions, recognition categories, etc.), and produces form registration data.

S330: The form registration data are registered in the form data table 32. The account number is registered in the account master table 31.

S340: The tables 32 and 31 are written into the file 22.

In this manner, a new form is registered. In this example, the search rules for the account number of the new form are the same as those of a form that has already been registered, so that no new search rules are registered. However, it is possible to register other search rules, when a form with different search rules is registered.

Other Embodiments

In addition to the above-described embodiments, the following variations of the present invention are also possible.

(1) In the foregoing, the account number was shown to be a 5 digits - 1 digit - 6 digits numerical character string, but it can also be another numerical character string. Moreover, it can also be a numerical character string without hyphens.

(2) The foregoing explanations related to a sheet feed-type, but the present invention can also be applied to an apparatus and method where the sheets are fixed.

Thus, the present invention attains the following effects.

The present invention takes advantage of the fact

that the payment form contains the payee account number, and that the form type is specified by this account number, so that the form is identified by recognizing the payee account number from the form image. Moreover, the account number is a character string of numerals and symbols, so that the account number can be identified on the form image with high precision by simple search rules (knowledge information). Therefore, the form can be identified accurately and fast. Furthermore, the payee account number is a portion of the input data of the payee form, so that form identification and data input can be performed simultaneously.

The invention has been described based on the embodiments. The invention, however, may be embodied in other specific forms without departing from the spirit or essential characteristics thereof, and these are not excluded from the scope of the invention.